Continuum estimation in low-resolution gamma-ray spectra based on deep learning*

Ri Zhao,[1,2,†] Li-Ye Liu,[1] Xin Liu,[1,2] Zhao-Xing Liu,[1,2] Run-Cheng Liang,[1,2] Ren-Jing Ling-Hu,[1] Jing Zhang,[1,2] and Fa-Guo Chen[1,2]

[1]*China Institute for Radiation Protection, Shanxi 030006, China*

[2]*Shanxi Provincial Key Laboratory for Translational Nuclear Medicine and Precision Protection, Shanxi 030006, China*

In this study, an end-to-end deep learning method is proposed to improve the accuracy of continuum estimation in low-resolution gamma-ray spectra. A novel process for generating the theoretical continuum of a simulated spectrum is established, and a convolutional neural network consisting of 51 layers and more than 10^5 parameters is constructed to directly predict the entire continuum from the extracted global spectrum features. For testing, an in-house NaI-type whole-body counter is used, and 10^6 training spectrum samples (20% of which are reserved for testing) are generated using Monte Carlo simulations. In addition, the existing fitting, step-type, and peak erosion methods are selected for comparison. The proposed method exhibits excellent performance, as evidenced by its activity error distribution and the smallest mean activity error of 1.5% among the evaluated methods Additionally, a validation experiment is performed using a whole-body counter to analyze a human physical phantom containing four radionuclides. The largest activity error of the proposed method is -5.1%, which is considerably smaller than those of the comparative methods, confirming the test results. The multiscale feature extraction and nonlinear relation modeling in the proposed method establish a novel approach for accurate and convenient continuum estimation in a low-resolution gamma-ray spectrum. Thus, the proposed method is promising for accurate quantitative radioactivity analysis in practical applications.

Keywords: Gamma-ray spectrum, continuum estimation, deep learning, convolutional neural network, end-to-end prediction.

I. INTRODUCTION

The continuum in a gamma-ray spectrum is typically defined as all the energy deposition counts in a detector, excluding photoelectric-effect events, and is thus formed by gamma-ray scattering [1]. As a baseline in a spectrum, the continuum should be estimated to extract the net counts of the photoelectric peak, which leads to radionuclide activity determination considering the detection efficiency. Therefore, continuum estimation in gamma-ray spectra is essential for quantitative radioactivity analysis [2]. However, accurate continuum estimation is often difficult using existing methods, particularly for low-resolution gamma-ray spectra, in which significant peak broadening increases the continuum complexity.

Three continuum estimation methods are available. 1) The fitting method is widely used in various applications [3–6]. It fits the peak region linearly or nonlinearly to a function with an added continuum. Given the difficulty in obtaining a precise characterization of the peak and continuum shapes and stabilizing the multiparameter fitting, deviations are likely for complex continuums. 2) The step-type method is adopted in commercial software packages such as Genie 2000 (Canberra Industries) and GammaVision (Ortec Industries) [7–9]. Because this method generates a step-shaped curve within a peak region that declines from left to right, it is suitable for a continuum exclusively formed by multiple Compton scattering events but not for one containing high background counts. 3) The topological method, called the peak erosion method, used in this study typically involves an iterative process to remove the peaks in convex structures and establish a baseline. Although numerous iterative processes have been proposed [10–17], the peak erosion method can roughly outline the continuum shape but fails to describe fine structures.

Overall, the three available methods perform approximations or have predetermined parameters and cannot suitably estimate complex continuums, particularly from low-resolution gamma-ray spectra. Moreover, they involve complicated data processing steps (e.g., fitting and erosion), which are inconvenient in practice. To improve the accuracy and applicability of continuum estimation, an end-to-end method based on deep learning is proposed and a novel approach compared to existing methods is established.

The remainder of this paper is organized as follows: Section II presents the proposed method for generating the theoretical continuum of a simulated gamma-ray spectrum, as well as a convolutional neural network (CNN) constructed to relate the primary gamma-ray spectrum to its continuum through deep learning. In addition, we describe test and validation experiments conducted using an in-house whole-body counter (WBC) to evaluate the proposed and three existing methods. The experimental results are reported in Section III, and the limitations of existing methods and advantages of the proposed method are further discussed in Section IV. Finally, conclusions are drawn in Section V.

II. MATERIALS AND METHODS

A. Generation of theoretical continuum

The measured gamma-ray spectrum is broadened owing to the statistical fluctuations of either light in the scintillation detector or electron-hole pairs in the semiconductor detector [18]. However, a spectrum without broadening can be

* This work was supported by the National Natural Science Foundation of China (grant number 12005198).

† Corresponding author, Ri Zhao, Xuefu Street 102, Taiyuan, Shanxi 030006, China, 15234152638, zhaor.abc@163.com

synthesized using Monte Carlo (MC) simulations. When a simulated spectrum without broadening is generated, each peak assumes a single channel, thereby simplifying the removal of net peak counts. The remaining spectrum can then be manually broadened, and the theoretical continuum of the corresponding broadened spectrum can be obtained. The theoretical continuum is obtained by the procedure illustrated in Fig. 1 and is described below.

Let $x = [x_1, x_2, \ldots, x_n]$ be a simulated gamma-ray spectrum and $x' = [x'_1, x'_2, \ldots, x'_n]$ be its corresponding spectrum without broadening, where n is the number of channels. If a peak exists at channel k in x', the net peak counts are removed by replacing the counts of channel k with the average counts of its adjacent left and right channels, as follows:

$$\begin{cases} x'_k = \frac{x'_{k-1}+x'_{k+1}}{2} & 1 < k < n \\ x'_k = x'_2 & k = 1 \\ x'_k = x'_{n-1} & k = n \end{cases} \quad (1)$$

If x' has multiple peaks, the above calculation is applied to each one. To obtain x^b according to the broadening function applied in the MC simulation, x' is manually broadened. Consider a common Gaussian broadening function given by

$$E_p = E_d + \sigma X_f \quad (2)$$

where E_d and E_p are the deposited energies before and after broadening during simulation, respectively; σ is the standard deviation of E_d; and X_f is a Gaussian random number. Manual broadening proceeds as follows [19].

$$x^b_i = \sum_{j=1}^{n} \frac{x'_j}{\sqrt{2\pi}\sigma w} e^{-\frac{(E_i - E_j)^2}{2\sigma^2}} \quad (3)$$

where x^b_i represents the counts of channel i in the broadened spectrum x^b; E_j and E_j are the energy values of channels i and j, respectively; and w is the channel width expressed in terms of energy. The resulting x^b corresponds to the theoretical continuum of the gamma-ray spectrum x.

B. CNN for continuum estimation

CNNs are the most common architecture used in deep learning [20–24]. Compared to a fully connected network, a CNN can extract multiscale features over multiple convolutional layers and prevent overfitting through parameter sharing when treating high-dimensional data in computer vision and other areas [25–27]. Considering a spectrum with hundreds or thousands of channels as the input and a predicted continuum over the entire spectral range as the output, high-dimensional data are involved at both ends. Hence, a CNN is preferable to a fully connected network for extracting distinctive and stable shape features from the spectrum and relating them to a continuum. Additionally, the convolution operation is highly effective for handling data with localized correlations or local information, as has been extensively demonstrated in image processing, where 2D local correlations are prevalent. Therefore, we propose the incorporation of CNN to harness their inherent ability to exploit the 1D local correlations exhibited in the spectra.

The proposed CNN is a modified version of ResNet-50 [28], the architecture of which is shown in Fig. 2. It consists of an input layer, multiple residual modules, a fully connected layer, and an output layer. The input spectrum is arranged in 1024 channels, a configuration commonly considered for a low-resolution detector. Spectra with different numbers of channels can be matched to the input via channel splitting or merging. Residual modules prevent the vanishing gradient problem during deep learning by establishing skip connections from the output of the front layer to the subsequent outputs across a convolutional layer [29, 30]. Two types of residual modules, denoted as R1 and R2, are used, and their architectures are shown in Fig. 2a and Fig. 2b, respectively. Module R1 is characterized by four parameters: R1($D_{in}, D_{out}, Ch_{in}, and Ch_{out}$), where D represents the data dimension, Ch represents the number of convolution kernels, and the subscripts in and out indicate the input and output, respectively. Module R2 is characterized by two parameters: R2($DandCh$), where D and Ch apply to both input and output data. In Fig. 2a and Fig. 2b, Conv1D(k, s) represents a 1D convolutional layer with a convolution kernel width k and stride s, and Ch_1, Ch_2, and Ch_3 represent the number of channels of the corresponding convolution kernels. In R1, $Ch_1 = Ch_2 = Ch_{out}/4$ and $Ch_3 = Ch_{out}$, and in R2, $Ch_1 = Ch_2 = Ch/4$ and $Ch_3 = Ch$. BN denotes the batch normalization applied to the batch training data, and ReLU denotes the rectified linear unit (ReLU) activation in each layer. ReLU activation is employed to introduce nonlinearity, thereby enhancing the mapping capability of the CNN and ensuring non-negativity in each channel of the continuum in the final layer. Each module R1 reduces the data dimensions by one-fourth and quadruples the number of convolution kernel channels (except for the first module), whereas R2 maintains the two parameters. Through four R1–R2 blocks, the spectral features are finally embedded into a 16×256 vector and mapped onto the continuum through the last fully connected layer. Because modules R1 and R2 have three convolutional layers, the entire CNN contains 51 layers, including the input and output layers, and more than 10^5 parameters.

C. Test setup

First, a laboratory test experiment was conducted using available equipment. The setup involved an in-house NaI-type WBC to measure the radioactivity from the human body. WBC is commonly used in occupational radiation monitoring in nuclear facilities (e.g., nuclear power plants) to determine the category and activity of radionuclides inside an exposed human body by detecting the emitted gamma-ray spectrum [31, 32].

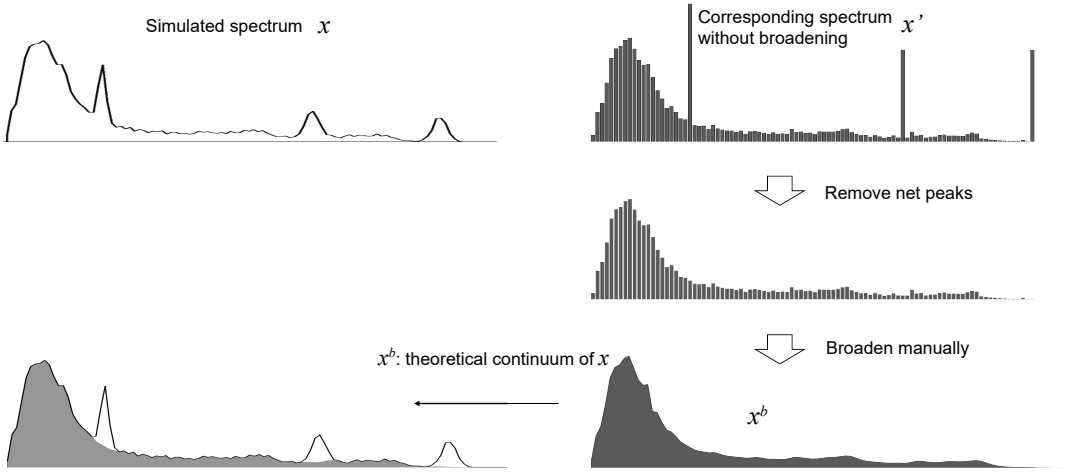

Fig. 1. Generation of theoretical continuum.

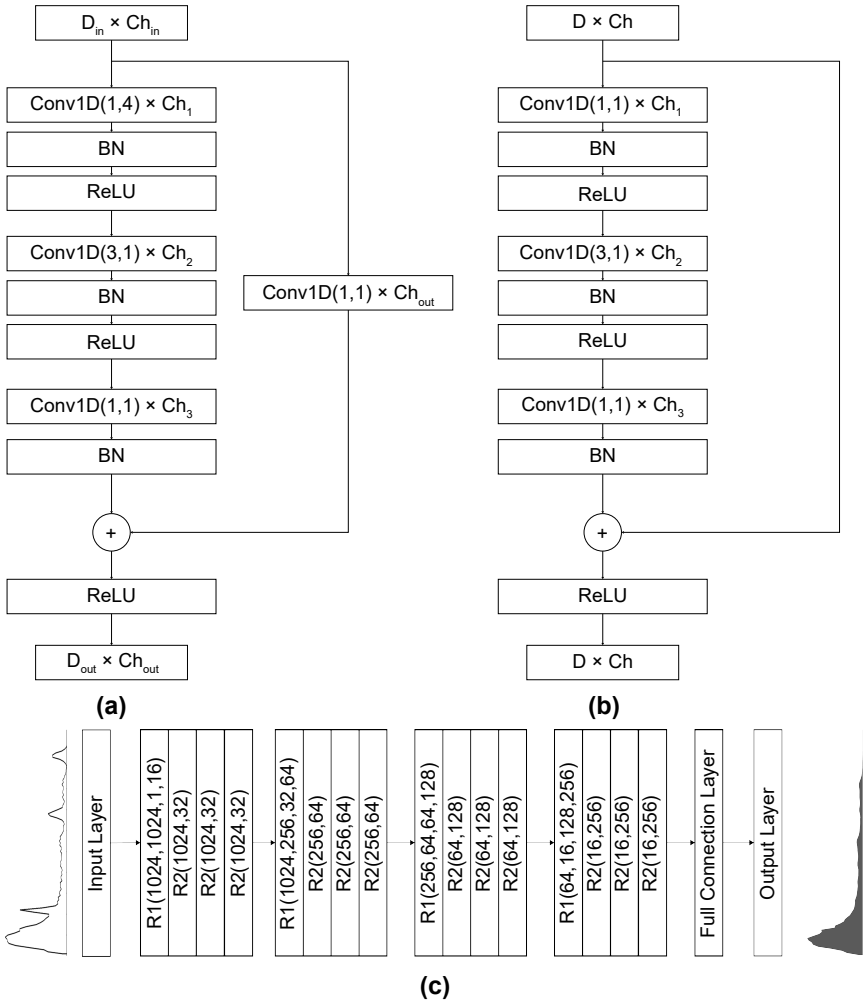

Fig. 2. Architectures of residual modules (a) R1, (b) R2, and (c) proposed CNN.

This setup is suitable for evaluating the continuum estimation for the following reasons: First, owing to the limited energy resolution of the NaI detector, the peaks in its spectrum are strongly broadened, leading to a wide continuum that is

more difficult to estimate accurately than narrow continuums. Second, the NaI detector used in the WBC is larger than that used in other devices. Specifically, the size of the WBC used in this study was $7.6\,\mathrm{cm} \times 12.7\,\mathrm{cm} \times 40.6\,\mathrm{cm}$, which resulted in multiple Compton scattering events inside the detector and a much higher continuum within the peak region than that obtained from a small detector. Clearly, a higher continuum requires better estimation to obtain accurate peak net counts. Third, the human body with radionuclides is a large volumetric radioactive source, and the emitted gamma rays are considerably scattered before detection. Consequently, more continuum counts are recorded in the low- and middle-energy regions of the spectrum. Meanwhile, the detector in a WBC is usually well-shielded by thick stainless steel and lead, causing severe backscattering of gamma rays. Combining the two abovementioned scattering mechanisms, the continuum differs completely from that observed when measuring a simple point source without shielding. Fourth, radionuclides in the human body have low activity; thus, the peak-to-continuum count ratios are lower than those formed by a strong source. This increases the importance of accurate continuum estimation.

D. Dataset construction

The digital model of the setup described in Section II C was constructed using a human body represented by a human phantom (see Fig. 3c). Accordingly, an MC simulation was conducted to generate a dataset for training and testing the CNN using the GEANT4 code [33–35].

Nine common radionuclides used in routine internal exposure monitoring in nuclear power plants were selected to simulate the spectra, as detailed in Table 1. Per simulation run, a random selection of one to five radionuclides was made from a given set. The activity of each radionuclide was then randomly assigned a value ranging from hundreds to thousands of becquerels, based on realistic activity levels. Simulated source particles were then included with energies equal to the gamma-ray theoretical energies and quantities equal to the radionuclide activities multiplied by the gamma-ray branching ratios for 5 min, which is a typical acquisition time in practice. In addition, the source particles were uniformly distributed in the phantom, which is consistent with real conditions. The real spectrum-broadening function in Eq. (4) was used during the simulation, and the gamma-ray energies shifted slightly from their theoretical values (as presented in Eq. (5)) to mimic the temperature shifts that occur in NaI detectors.

$$\sigma = -3.46 + 0.98\sqrt{E_d} \tag{4}$$

where σ and E_d are both expressed in keV according to the definitions provided in Eq. (2).

$$E_\gamma^s = E_\gamma + 0.022 E_\gamma \xi \tag{5}$$

where E_γ and E_γ^s are the gamma-ray energies before and after shifting, respectively; ξ is a random number between -1 and 1; and the value of 0.022 is also determined through measurements.

We obtained a training set with 10^6 spectrum samples, 20% of which were reserved as test samples. The corresponding continuums were generated using the method described in Section II A.

Table 1. Nine common radionuclides in occupational internal exposure monitoring of nuclear facilities and their radiation information.

Radionuclide	Energy [keV] (branch ratio, %)
^{60}Co	1173.2 (100.0), 1332.5 (100.0)
^{137}Cs	661.7 (85.1)
^{134}Cs	567.0 (23.8), 604.7 (97.6), 797.0 (94.1)[a]
^{57}Co	122.1 (85.5)
^{59}Fe	1099.2 (56.5) 1292.6 (43.2)
^{54}Mn	834.8 (100.0)
^{51}Cr	320.1 (9.8)
^{65}Zn	1115.5 (50.8)
^{95}Nb	765.8 (99.8)

[a] 563.2 keV (8.4%) and 569.3 keV (15.4%) gamma rays are merged into 567.0 keV (23.8%) owing to their similar energy. Likewise, 795.8 keV (85.4%) and 801.9 keV (8.7%) gamma rays are merged into 797.0 keV (94.1%).

E. Evaluation measures

Instead of directly assessing the error in the estimated continuum counts, the proposed method was evaluated more intuitively by comparing the deduced radionuclide activity values with the theoretical values defined during the simulation. The activity relative error (AcE) and mean AcE (MAcE) obtained from the test set were used for evaluation.

Consider a peak region that includes n channels: The activities of the radionuclides were determined as follows:

$$A = \frac{S}{\varepsilon T \eta} = \frac{\sum_{i=1}^{n} (Y_i - C_i)}{\varepsilon T \eta} \tag{6}$$

where S is the sum of the peak net counts; Y_i and C_i are the total counts and estimated continuum counts in channel i, respectively; n is the number of channels in the peak region; ε is the photoelectric efficiency determined by the MC simulation; T is the acquisition time (5 min for a test spectrum); and η is the branch ratio. For a multiplet, S is determined per peak using nonlinear least-squares fitting, similar to the fitting method for continuum estimation. When multiple peaks are involved for one radionuclide, its activity is given by the weighted average of the activity across the peaks, as demonstrated below.

$$\bar{A} = \frac{\sum_{p=1}^{P} \frac{A_p}{\sigma_{A_p}^2}}{\sum_{p=1}^{P} \frac{1}{\sigma_{A_p}^2}} \tag{7}$$

where A_p and $\sigma^2_{A_p}$ are the activity and its uncertainty estimated based on peak p, respectively, and P is the total number of peaks of this radionuclide.

$\sigma^2_{A_p}$ is determined by the error propagation based on Eq. (7). Because of the challenge of accurately evaluating the relative error of C_i, the Poisson distribution was utilized as an approximation to simplify the distribution of C_i. Consequently, $\sigma^2_{A_p}$ can be expressed as follows:

$$\sigma^2_{A_p} = \frac{\sum_{i=1}^{n}(Y_i + C_i)}{\varepsilon T \eta} \tag{8}$$

Based on the activity of each radionuclide, the $\mathrm{AcE_j}$ and MAcE are defined as follows:

$$AcE_j = \frac{\hat{A}_j - A_j}{A_j} \tag{9}$$

$$MAcE = \sum_{j=1}^{m} \frac{\left|\hat{A}_j - A_j\right|}{mA_j} \tag{10}$$

where $\hat{A}_j$ and A_j are the estimated and theoretical activity values of radionuclide j, respectively; and m is the number of radionuclides in the test spectrum samples.

F. Comparison methods

To demonstrate the high performance of the proposed method, we compared it with three existing continuum estimation methods—the fitting, step-type, and peak erosion methods—on the same training and test sets.

The fitting and step-type methods are only applicable to the peak regions. Thus, performance evaluation was limited to the peak regions of each spectrum. The peak region was defined as the range from left to right of the peak centroid with a full width at half maximum of 1.5. Overlapping peak regions formed by adjacent peaks were treated as a single region.

Details of the comparison methods can be found in corresponding studies, and we provide brief descriptions for convenience.

1. *Fitting method*

Datapoints in the peak region can be fitted by the peak function P added to the continuum function C. To this end, weighted nonlinear least-squares fitting was applied to determine the minimum value of the following function:

$$L_\theta = \sum_i w_i \left(Y_i - P_i - C_i\right)^2 \tag{11}$$

where Y_i, P_i, and C_i are the total counts, peak net counts, and continuum counts in channel i, respectively; w_i is the channel weight, which is set to $1/Y_i$ assuming a Poisson distribution; and θ is the parameter to be optimized.

Considering the spectrum acquired by the NaI detector, the following Gaussian function for a singlet can be used:

$$P_i = H_P e^{-\frac{(i-c)^2}{2\sigma^2}} \tag{12}$$

where H_P is the peak amplitude, i is the channel index, c is the peak centroid, and σ is a parameter related to the peak width, which is given by

$$\sigma = \frac{FWHM}{2.355} \tag{13}$$

where $FWHM$ is the full width at half maximum. For a multiplet, the sum of singlet functions should be applied.

The continuum function C has several representations. Theoretically, the complementary error function, that is, the convolution of a Gaussian function with a negative step function centered at the peak centroid, allows correct estimation of multiple Compton scattering counts in the continuum [36, 37], and extra background counts can be accounted for by adding a linear term. Thus, an ideal C is given by

$$C_i = H_C \,\mathrm{erfc}\left(\frac{i-c}{\sqrt{2}\sigma}\right) + ai + b \tag{14}$$

where H_C is the function amplitude; i is the channel index; c is the peak centroid; σ is a parameter related to the peak width; a and b are linear parameters; and $erfc$ is the following complementary error function:

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt \tag{15}$$

Therefore, θ embeds H_P, H_C, c, σ, a, and b.

Several experiments have shown that C given by Eq. (14) is highly complex, and unsupervised fitting easily fails for multiplets in a low-resolution spectrum, thereby providing meaningless results. To ensure a suitable solution, we considered C as a simple cubic polynomial given by

$$C_i = a_1 i^3 + a_2 i^2 + a_3 i + a_4 \tag{16}$$

In addition, we used the Levenberg-Marquardt algorithm [38] to optimize Eq. (11).

2. *Step-type method*

The step-type method implementations in Genie 2000 and GammaVision differ but provide similar results [8, 9]. The implementation of Genie 2000 is based on a direct and brief formula that is simpler and clearer than the iterative process used in GammaVision. Thus, we selected the implementation in Genie 2000, which is formulated as follows:

$$C_i = \frac{C_1}{n} + \frac{C_2 - C_1}{nG} \sum_{j=1}^{i} Y_j \quad (17)$$

where G is the total sum of counts (gross) in the peak region; n is the number of continuum channels on each side of the region; C_1 and C_2 are the sums of counts in the continuum region to the left and right of the peak, respectively; and Y_j is the total count in channel j. In Eq. (17), the first derivative (i.e., first discrete difference) of the continuum in a channel is assumed to be inversely proportional to its total count [39, 40]. Hence, the estimated continuum declines from left to right across the peak region and exhibits an obvious step shape near the peak centroid, which explains the name of this method.

3. *Peak erosion method*

Although various iterative erosion methods have been proposed, we used its simplest version [10, 15], which can be described by pseudocode as follows:

For j = 1 : M
  For i = 1 : N
    Y_i = min(Y_i, $\frac{Y_L+Y_R}{2}$)
  End
End

Here, M is the number of iterations; N is the number of channels; Y_L and Y_R are the counts in the channels to the left and right of channel i for 1.5 of the full width at half maximum, respectively; and min is the minimum function. After erosion, Y is a continuum across the entire spectrum. We selected $M = 8$ as the optimal value.

G. Validation experiment

After testing, a validation experiment was conducted by measuring a human physical phantom using WBC (Fig. 3). The in-house phantom contained uniformly distributed ^{134}Cs, ^{137}Cs, ^{57}Co, and ^{60}Co with the known activity listed in Table 2. We performed 100 repeated measurements, and the average activity of each radionuclide estimated by the proposed method was validated by comparing it with the true value and the results of the three comparison methods. The required detection efficiency for each gamma ray was determined in the same manner as in the test step.

Table 2. Radionuclides and their activities inside human physical phantom.

Radionuclide	^{57}Co	^{134}Cs	^{137}Cs	^{60}Co
Activity (Bq)	5498.2	3849.6	2879.5	4023.1

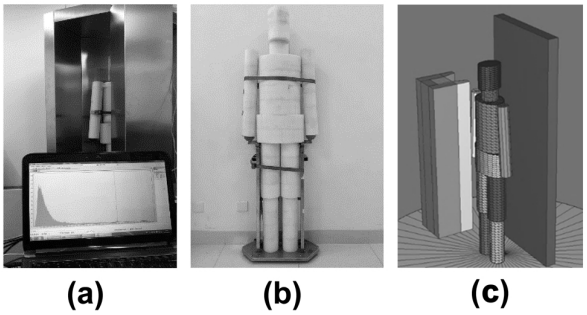

Fig. 3. In-house (a) WBC, (b) human physical phantom, and (c) corresponding digital model constructed in GEANT4.

III. RESULTS

A. Test

The AcE distribution and MAcE of each method are shown in Fig. 4 and Table 3. The AcE of the proposed method is within $\pm 3\%$ for all test samples, leading to the smallest MAcE of 1.5%. The fitting method provides a relatively small AcE of -6%–10% under most conditions but shows some outliers up to -40%–60%, resulting in an MAcE of 5.5%. The step-type method provided an AcE within $\pm 8\%$, achieving the second-best results among all the methods with an MAcE of 3.1%. The peak erosion method had the worst performance, with its AcE ranging from -20% to 90%, resulting in the largest MAcE of 18.2%.

Table 3. MAcE of evaluated methods.

	Proposed CNN	Fit	Step	Erosion
MAcE (%)	1.5	5.5	3.1	18.2

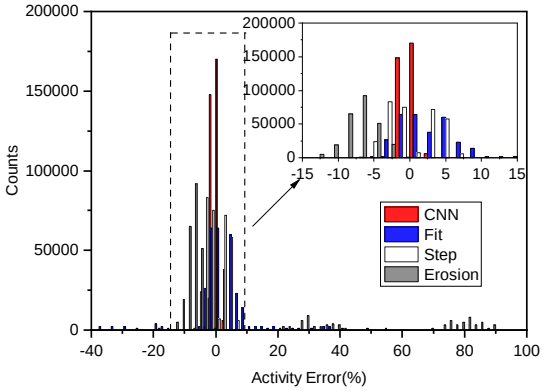

Fig. 4. AcE distribution of proposed (CNN), fitting (Fit), step-type (Step), and peak erosion (Erosion) methods.

B. Typical test scenarios

During testing, the performance of the compared methods was evaluated for three typical scenarios (see Fig. 5).

1. *Singlet without interference*

The singlet of ^{137}Cs and its continuum estimated using each method are shown in Fig. 5a and Fig. 5b. The primary spectrum (Spec), theoretical continuum (TC), and results of the proposed CNN-based (CNN), fitting (Fit), step-type (Step), and peak erosion (Erosion) methods are demonstrated. Moreover, the goodness of fit was evaluated using the coefficient of determination, R^2, by constructing datapoints $(C_i, \hat{C}_i)$ in the coordinates of the theoretical (C) and estimated $(\hat{C})$ continuum counts (see Fig. 5c). The coefficient R^2 is calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(C_i - \hat{C}_i\right)^2}{\sum_{i=1}^{n} \left(C_i - \bar{C}\right)^2} \tag{18}$$

where n is the number of continuum channels and $\bar{C}$ is the mean theoretical continuum count. We use R^2 because it can completely describe the closeness between the estimated and theoretical continuums while avoiding division by zero.

Other quantities are more intuitive and direct but are not suitable for evaluation. For instance, the relative error of the continuum counts per channel is defined as

$$RE_i = \frac{\left(\hat{C}_i - C_i\right)^2}{C_i} \tag{19}$$

In addition, the mean count error is defined as

$$MCE = \sum_{i=1}^{n} \frac{\left(\hat{C}_i - C_i\right)^2}{nC_i} \tag{20}$$

However, the relative error is highly sensitive to small C_i values owing to its denominator, which consequently skews the assessment. Similarly, the mean count error tends to make small C_i dominant. Moreover, if C_i approaches zero, these values will reach infinity.

The proposed method achieved a coefficient R^2 of 0.9998, indicating a nearly ideal estimation, followed closely by the step-type method with R^2 of 0.9992, the fitting method with a smaller R^2 of 0.9704, and the peak erosion method with the lowest R^2 of 0.8994, showing an obvious deviation in the estimation.

2. *Singlet on high background counts*

A singlet with a high background count was used to establish different scenarios. In Fig. 5d, a 661.7 keV gamma ray of ^{137}Cs and 1099.2 keV and 1292.6 keV gamma rays of ^{59}Fe increase the counts in the low-energy spectrum region, thus changing the continuum shape under the singlet of ^{57}Co. The proposed method exhibited the highest performance with R^2 of 0.9992, followed by the fitting method with R^2 of 0.9952, whereas the step-type method demonstrated a low performance with R^2 of 0.9753. The results of the peak erosion method are not shown in Fig. 5f because its error is excessively high, providing an opposite trend with an R^2 value of -0.9201.

3. *Overlapping multiplet*

A more complex scenario is illustrated in Fig. 5g. Three peaks of ^{134}Cs and the peak of ^{137}Cs highly overlap, resulting in a more complex continuum compared to the scenarios reported in Sections 3.2.1 and 3.2.2. The R^2 value of the proposed method was 0.9964, outperforming the step-type, fitting, and peak erosion methods with R^2 values of 0.9890, 0.9720, and 0.8755, respectively.

C. Validation

The results obtained from the measurements listed in Table 4 show the largest AcE of -5.1%, -55.1%, -8.3%, and 99.1% for the proposed, fitting, step-type, and peak erosion methods, respectively. Similar to testing results, the proposed method provides the best estimation, whereas the step-type method yields the second-best results with AcE of less than 10% for the four radionuclides, and the fitting method provides the third-best results with good estimation for radionuclides ^{134}Cs, ^{137}Cs, and ^{60}Co but poor estimation for ^{57}Co. Additionally, the peak erosion method again yields the worst results with its high AcE of 99.1% for ^{57}Co.

Table 4. Activity estimation of evaluated methods on measured spectrum.

Nuclide	^{57}Co	^{134}Cs	^{137}Cs	^{60}Co
CNN [Bq] (error, %)	5718.1 (4.0)	3726.4(-3.2)	2732.6 (-5.1)	4180 (3.9)
Fit [Bq] (error, %)	2468.7(-55.1)	4049.8 (5.2)	3032.1 (5.3)	3821.9 (-5.0)
Step [Bq] (error [%])	5866.6 (6.7)	3661 (-4.9)	2640.5 (-8.3)	3721.4 (-7.5)
Erosion [Bq] (error, %)	10946.9 (99.1)	3568.6 (-7.3)	2703.9 (-6.1)	3765.6 (-6.4)

IV. DISCUSSION

A. Test results

The AcE distributions and MAcE values reported in Section III B can be explained by the limitations of existing methods and the advantages of the proposed method.

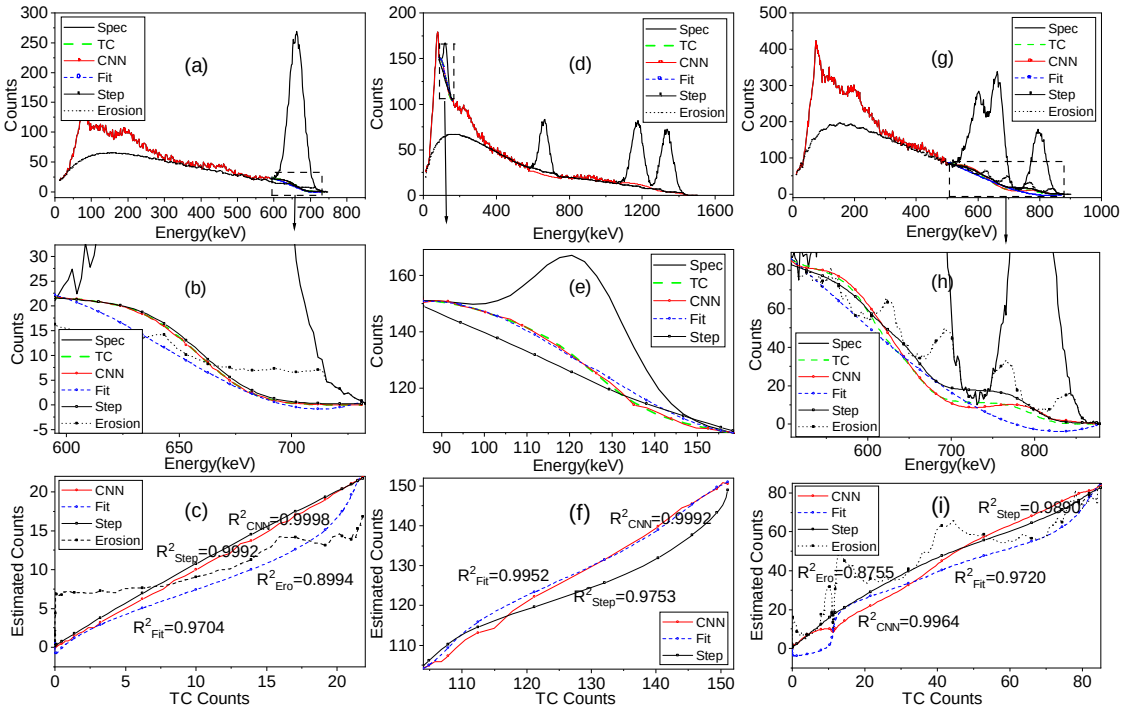

Fig. 5. Estimation results of each method on singlet of ^{137}Cs ((a) full estimation, (b) magnified view, (c) theoretical vs. estimated counts), on singlet of ^{57}Co ((d) full estimation, (e) magnified view, (f) theoretical vs. estimated counts), and on multiplet of ^{134}Cs and ^{137}Cs ((g) full estimation, (h) magnified view, (i) theoretical vs. estimated counts).

The fitting method involves multiparameter optimization, which is highly nonlinear, nonconvex, and sensitive to parameter initialization. Hence, this method can easily fail for a wide and complex continuum when performed automatically without manual adjustments, resulting in considerable errors as shown in Fig. 4. In addition, the estimation is determined by a fitting function that limits the representable continuums. Overall, the fitting method is unstable for continuum estimation in low-resolution gamma-ray spectra because its unpredictable results may be suitable under simple conditions but unacceptable for high and complex continuums.

The step-type method can describe continuum counts within a peak region formed by multiple Compton scattering events of the concerned gamma ray but cannot determine the background counts. The detailed theory of this method can be found in existing literature [40, 41]. This method can only provide a continuum decline from left to right across the peak region and correct results when the background counts are negligible. However, it deviates when the continuum shape changes significantly from the ideal step curve for low-energy singlets or complex multiplets.

The peak erosion method considers the convexity of the peak structure and generates a relatively flat curve across the entire spectrum. However, when the continuum is convex, a large estimation error is observed. Moreover, the generated curve exhibits a random shape and cannot represent the details of a real continuum. Consequently, this method exhibits the worst performance in most scenarios.

Unlike existing methods, the proposed method estimates the continuum using global spectrum features extracted by a CNN, which provides small-scale count variance and other statistical characteristics, as well as large-scale count correlation and shape characteristics over the entire spectrum. Thus, it outperforms the fitting and step-type methods, which use limited local counts within peak regions, and the peak erosion method, which uses counts on each side of the concerned channel within 1.5 times the full width at half maximum. In fact, over a spectrum, the continuum of a local region is highly related to the counts in other regions; however, this relationship is too complex and nonlinear to be modeled by conventional methods. In contrast, a high-performance CNN is suitable for complex nonlinear mapping. By linking the primary spectrum to its continuum via multiple convolutional layers, the photoelectric peak, Compton scattering content, background radiation, backscattering counts, and other components are integrated by the CNN for prediction, establishing an end-to-end continuum estimation without any explicit regression or additional data processing. Therefore, the proposed method is convenient and accurate.

B. Analysis of selected scenarios

The limitations of the existing methods and the advantages of the proposed method can be further demonstrated by considering the scenarios detailed in Section III B.

1. Singlet of ^{137}Cs without interference

For a singlet without interference from other high-energy rays, if the environmental background counts are subtracted, the continuum in the singlet is exclusively formed by Compton scattering of the concerned gamma ray and shows an ideal step shape. Thus, the step-type method provides high accuracy in this scenario, with an R^2 of 0.9992. However, this value is still lower than that of the proposed method, demonstrating the advantage of global spectrum feature extraction over local continuum estimation. The bias of the fitting method is also relatively small, indicating a suitable step-shape estimation using a cubic polynomial. Nevertheless, owing to the inherent shape of a cubic polynomial and the lack of non-negative control during fitting, the continuum counts estimated by the fitting method are negative in the region exceeding 700 keV, leading to meaningless results. Nevertheless, this estimation can be applied to calculate subsequent peak net counts. Despite its simple continuum shape, the peak erosion method achieves the lowest R^2 value because of its estimation irregularity. Interestingly, the sum of the estimated continuum counts may exhibit less deviation than the continuum shape, as shown in Fig. 5b, where the continuum is first underestimated and then overestimated.

2. Singlet of ^{57}Co on high-energy background

The simulated spectrum for the scenario reported in Section III B 2 without ^{57}Co is shown in Fig. 6a and Fig. 6b. The counts from the scattered gamma rays of 661.7 keV, 1099.2 keV, and 1292.6 keV form the background of the peak of ^{57}Co at $85.3-158.7$ keV. When added to the original step-type continuum formed by multiple Compton scatterings of a 122.1 keV gamma ray, the additional background counts substantially change the final continuum shape. In Fig. 6 (b), the background curve fluctuates at $85.3-120$ keV and drops sharply afterward. Thus, the continuum of ^{57}Co initially experiences a slower decline than expected, followed by a rapid decline. This indicates poor performance of the step-type method, as shown in Fig. 5f, which first underestimates and then overestimates the continuum.

Owing to least-squares optimization, the fitting method can adjust its shape more flexibly than the step-type method and thus displays a better result under this scenario when the correct fitting is obtained. However, a large error in the peak erosion method is evident.

Existing methods fail to predict the background of ^{57}Co contributed by ^{137}Cs and ^{59}Fe, thereby limiting their performance when multiple radionuclides are involved. In contrast, the proposed method bridges the counts in different energy regions through deep learning by relying on training samples and can implicitly estimate the background curve based on the peaks of ^{137}Cs and ^{59}Fe, as well as other spectrum characteristics, resulting in the best estimation.

3. Overlapping multiplet of ^{137}Cs and ^{134}Cs

The complexity of the continuum in the overlapping multiplet region from 487.9 keV to 888.6 keV clearly shows in the corresponding spectrum without broadening in Fig. 6c. The Compton edges of the 604 keV, 661 keV, and 795 keV gamma rays accumulate in the multiplet region, leading to the multistep continuum shown in Fig. 6d, which declines more slowly than the normal step shape in the region below approximately 600 keV but faster in $600-800$ keV. This trend agrees with the error of the step-type method under the same circumstances. Moreover, the estimated continuum of the step-type method is above the spectrum around the minimum spectrum value between the peaks at 661.7 keV and 797.0 keV, resulting in negative peak net counts in that region. However, the fitting method also fails to reflect the multistep shape (Fig. 5i) through a cubic polynomial, and negative continuum counts are again observed in the region around 800 keV in Fig. 5h. By contrast, the proposed method provides a higher estimation performance owing to the use of the CNN, even for the considered complex multiplet.

C. Validation results

Figure 7 shows the continuum estimated by each method for the singlet of ^{57}Co ($102.1-151.9$ keV), multiplet of ^{134}Cs and ^{137}Cs ($502.6-873.3$ keV), and multiplet of ^{60}Co ($1080.4-1430.7$ keV). The theoretical continuum is not shown in Fig. 7 because it is not visible in the measured spectrum, as explained in Section II A.

For the singlet, the performances of the proposed, step-type, and peak erosion methods are similar to those reported in Section III B 2. However, the fitting method performs significantly worse (Fig. 7c), possibly due to incorrect fitting, leading to a 55.1% underestimation of the activity of ^{57}Co. Moreover, the AcE of the proposed method (Table 4) is slightly higher than its MAcE (Table 3) because of the additional error induced by the difference between the simulated detection efficiency and the true value, as in the step-type method.

The results for the two multiplets agree with those reported in III B 3. The overestimation of the step-type method near the minima between the two overlapping peaks (approximately 725 keV and 1250 keV) in Fig. 7b and Fig. 7d and the negative estimation of the fitting method above 1350 keV in Fig. 7d are also observed.

V. CONCLUSION

Continuum estimation of the gamma-ray spectra is essential for assessing radionuclide activity. However, the invisibility and complexity of the continuum hinder accurate estimations, particularly for low-resolution spectra. Existing methods, including fitting, step-type, and peak erosion methods, have inherent limitations in terms of accuracy and applicability owing to their processing steps. To improve continuum

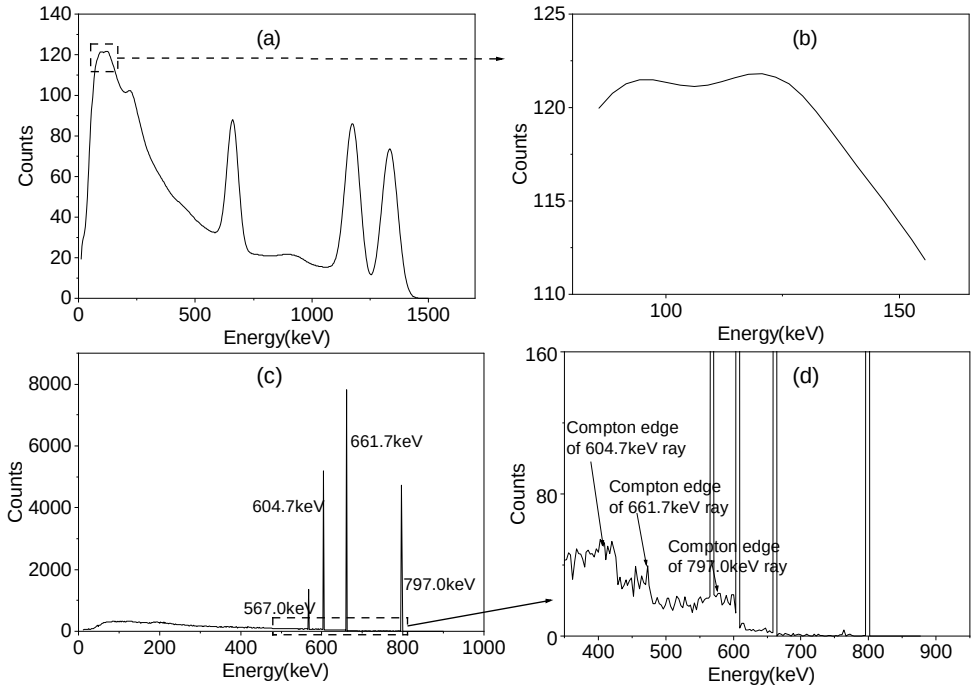

Fig. 6. Simulated spectrum of Fig. 5d without ^{57}Co: ((a) full and (b) magnified views) of Fig. 5g with broadening and ((c) full and (d) magnified views) without broadening.

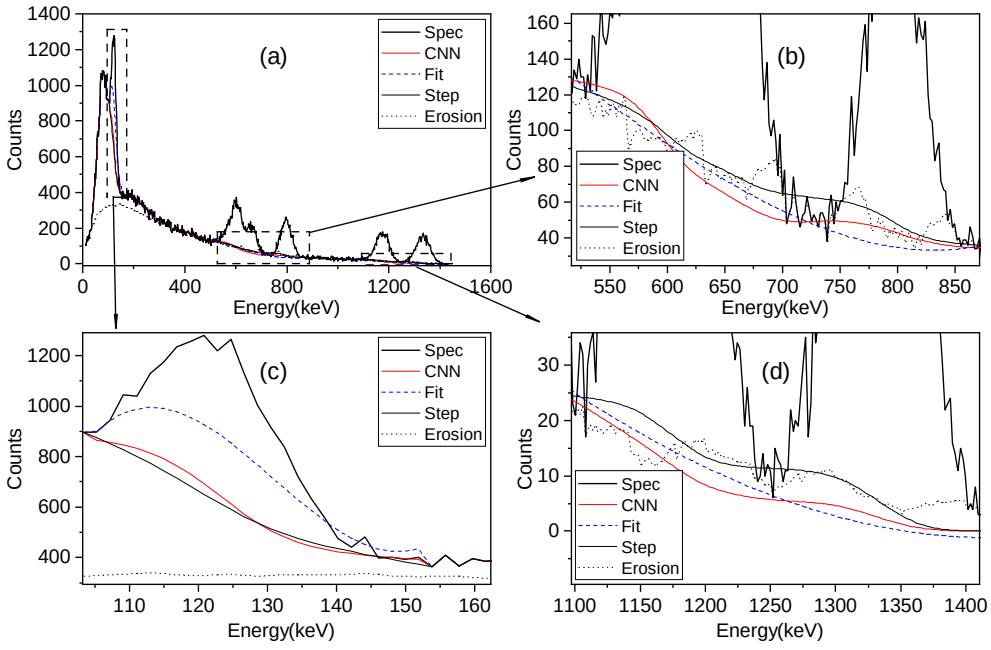

Fig. 7. Estimation results of each method on measured spectrum. (a) Full estimation and magnified views for (b) multiplet of ^{134}Cs and ^{137}Cs, (c) singlet of ^{57}Co, and (d) multiplet of ^{60}Co.

estimation, an end-to-end method based on deep learning was proposed in this study. The theoretical continuums of simulated spectra were generated as the ground truth for learning, and a CNN architecture with four R1-R2 blocks was used to determine the relationship between the primary spectrum and its corresponding continuum through training. The trained

CNN directly predicts the entire continuum across all channels. To test this method, a laboratory experiment was performed using an in-house WBC and 10^6 training spectrum samples generated through MC simulation. The test results showcased the superior performance of the proposed method, as indicated by its best AcE distribution and the smallest MAcE of 1.5% among the evaluated methods. Three typical testing scenarios were selected for further analysis. The proposed method performed global feature extraction for local continuum prediction, achieving an R^2 value that is closest to one across all scenarios. By contrast, the fitting method showed deviations for a complex continuum, whereas the step-type method initially underestimated and then overestimated the continuum for singlets containing high background counts, and reversed its performance for multiplets. The peak erosion method exhibited the worst performance owing to its rough estimation. Moreover, negative continuum counts occurred for the fitting and step-type methods. The results of a validation experiment using in-house WBC to measure a human phantom containing four types of radionuclides were consistent with the test results.

As an added benefit, the proposed method facilitated peak identification, particularly for weak and overlapping peaks. This was achieved by estimating a baseline across the entire spectrum, unlike the fitting and step-type methods that required local peak identification, thus enhancing applicability.

Overall, the proposed method provided accurate and convenient continuum estimation in a low-resolution gamma-ray spectrum, which can potentially enhance the accuracy of quantitative radioactivity analysis.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no competing financial interests or personal relationships that may have influenced the work reported in this study.

REFERENCES

[1] Q.X. Zhang, The Character of Airborne Gamma-ray Spectrometry and The Method for Spectrum Analysis, Chengdu University of Technology, 2011.

[2] X.Y. Ai, Y.X. Wei, W.Y. Xiao, Direct demodulation method for γ spectra analysis detected by CZT detectors. Journal of Tsinghua University (Science and Technology), 821–824 (2006). doi:10.16511/j.cnki.qhdxxb.2006.06.018

[3] J. Pang, γ spectrum data analysis. Shaanxi Science And Technology Press, XiAn, 1990.

[4] R. Qi, Study and Implementation of A Software used in Fast NaI Whole Body Counter γ Spectrum Analysis, LanZhou University, 2008.

[5] X.Y. Ai, Research on the Gamma Ray Spectral Response Characteristics and Relevant Spectral Analysis Method of Cadmium Zinc Telluride (CdZnTe) Detector, Tsinghua University, 2005.

[6] L Tang, X.K. Ma, K.B. Shi et al., A method for correcting characteristic X-ray net peak count from drifted shadow peak. Nucl. Sci. Tech. **34**, 175 (2023). doi:10.1007/s41365-023-01333-w

[7] A. Luca, J. Morel, Influence of the background approximation methods on the analysis of gamma-ray spectra. Appl. Radiat. Isotopes. **60**, 233–237 (2004). doi:10.1016/j.apradiso.2003.11.023

[8] Canberra Industries, Genie 2000 Customization Tools Manual, Canberra Industries, Meriden, 2006.300–310.

[9] Advanced Measurement Technology, GammaVision V6 User's Manual, Advanced Measurement Technology, Inc., Oak Ridge, 2006.180–185.

[10] L.V. East, R.L. Phillips, A.R. Strong, A fresh approach to NaI scintillation detector spectrum analysis. Nucl. Instrum. Methods. **193**, 147–155 (1982). doi:10.1016/0029-554X(82)90689-9

[11] A. Brunetti, Removal of the continuum of X-ray spectra using morphological operators. IEEE. Trans. Nucl. Sci. **45**, 2281–2287 (1998). doi:10.1109/23.725265

[12] N. Kourkoumelis, Continuum determination in spectroscopic data by means of topological concepts and Fourier filtering. Nucl. Instrum. Meth. A **691**, 1–4 (2012). doi:10.1016/j.nima.2012.06.061

[13] M.H. Zhu, L.G. Liu, Y.S. Cheng et al., Iterative estimation of the background in noisy spectroscopic data. Nucl. Instrum. Meth. A **602**, 597–599 (2009). doi:10.1016/j.nima.2009.01.174

[14] Q.X. Zhang, L.Q. Ge, Y. Gu et al., Background estimation based on Fourier Transform in the energy-dispersive X-ray fluorescence analysis. X-Ray Spectrom. **41**, 75–79 (2012). doi:10.1002/xrs.2360

[15] M. Morhác, An algorithm for determination of peak regions and baseline elimination in spectroscopic data. Nucl. Instrum. Meth. A **600**, 478–487 (2009). doi:10.1016/j.nima.2008.11.132

[16] W.M. Yin, H.Z. Liu, B. Tang, Discussion and Application of Eliminating the Background in γ-ray Spectrum by SNIP Algorithm. Journal of East China Institute of Technology. **32**, 245–248 (2009). doi:10.3969/j.issn.1674-3504.2009.03.010

[17] H. Yang, X.Y. Zhang, W.G. Gu et al., A novel method for gamma spectrum analysis of low-level and intermediate-level radioactive waste. Nucl. Sci. Tech. **34**, 87 (2023). doi:10.1007/s41365-023-01236-w

[18] H. Yu, M.M. Zhang, Y.Y. Du et al., Analysis on Energy Spectra for CdZnTe Gamma Ray Detector. Journal of Synthetic Crystals. **50**, 1884–1891 (2021). doi:10.16553/j.cnki.issn1000-985x.20210908.001

[19] W. Tang, J.G. Liang, Y. Ge et al., A method for neutron-induced gamma spectra decomposition analysis based on Geant4 simulation. Nucl. Sci. Tech. **33**, 154 (2022). doi:10.1007/s41365-022-01144-5

[20] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks. Commun. ACM **60**, 84–90 (2017). doi:10.1145/3065386

[21] H.J. Yoo, Deep Convolution Neural Networks in Computer Vision: A Review. IEIE. Trans. Smart. Process. Comput. **4**, 35–43 (2015). doi:10.5573/IEIESPC.2015.4.1.035

[22] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning. MIT Press, Cambridge, 2016.

[23] H.L. Liu, H.B. Ji, J.M. Zhang et al., A novel approach for feature extraction from a gamma-ray energy spectrum based on image descriptor transferring for radionuclide identification. Nucl. Sci. Tech. **33**, 158 (2022). doi:10.1007/s41365-022-01150-7

[24] F. Li, X.F. Huang, Y.L. Chen et al., Quantitative algorithm for airborne gamma spectrum of large sample based on improved shuffled frog leaping–particle swarm optimization convolutional neural network. Nucl. Sci. Tech. **34**, 112 (2023). doi:10.1007/s41365-023-01265-5

[25] F.Y. Zhou, L.P. JIN, J. Dong, Review of Convolutional Neural Network. Chinese Journal of Computer. **40**, 1230–1251 (2017). doi:10.11897/SP.J.1016.2017.01229

[26] Z.J. Sun, L. Xue, Y.M. Xu et al., Overview of deep learning. Application Research of Computers. **29**, 2806-2810 (2012), doi:10.3969/j.issn.1001-3695.2012.08.002

[27] L. Chang, X.M. Deng, M.Q. Zhou et al., Convolutional Neural Networks in Image Understanding. Acta. Automatica. Sinica. **42**, 1300–1312 (2016) doi:10.16383/j.aas.2016.c150800

[28] K.M. He, X.Y. Zhang, S.Q. Ren et al., Deep Residual Learning for Image Recognition. IEEE Conf. Comput. Vis. Pattern Recognit. 2016. doi:10.1109/CVPR.2016.90

[29] K. Zhang, X.H. Feng, Y.R. Guo et al., Overview of deep convolutional neural networks for image classification. Journal of Image and Graphics. **26**, 2305–2325 (2021). doi:10.11834/jig.200302

[30] Z.P. Zhao, Z Luo, P.Y. Wang et al., Survey on Image Classification Algorithms Based on Deep Residual Network. Computer Systems & Applications. **29**, 14–21 (2020). doi:10.15888/j.cnki.csa.007243

[31] R. Zhao, L.Y. Liu, Q.J. Cao, γ Spectrum Analysis Based on Reconstruction Technology for Whole Body Counter. At. Energy Sci. Technol. **53**, 1495–1501 (2019). doi:10.7538/yzk.2018.youxian.0733

[32] R. Zhao, N. Liu, Gamma spectrum analysis method based on deep learning model. At. Energy Sci. Technol. **57**, 379—388 (2023). doi:10.7538/yzk.2022.youxian.0215

[33] S. Agostinelli, J. Allison, K. Amako et al., Geant4-a simulation toolkit. Nucl. Instrum. Meth. A **506**, 250–303 (2003). doi:10.1016/S0168-9002(03)01368-8

[34] J. Allison, K. Amako, J. Apostolakis et al., Geant4 Developments and Applications. IEEE. Trans. Nucl. Sci. **53**, 270–278 (2006). doi:10.1109/TNS.2006.869826

[35] J. Allison, K. Amako, J. Apostolakis et al., Recent developments in Geant4. Nucl. Instrum. Meth. A **835**, 186–225 (2016). doi:10.1016/j.nima.2016.06.125

[36] A. Simonits, J. Östör, S. Kalvin et al., HyperLab: A new concept in gamma-ray spectrum analysis. J. Radiational. Nucl. Chem. **257**, 589–595 (2003). doi:10.1023/A:1025400917620

[37] J. Uher, G. Roach, J. Tickner, Peak fitting and identification software library for high resolution gamma-ray spectra. Nucl. Instrum. Meth. A **619**, 457–459 (2010). doi:10.1016/j.nima.2009.12.086

[38] D.W. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J. Soc. Ind. Appl. Math. **11**, 431–441 (1963). doi:10.1137/0111030

[39] R. Zhao, Research on γ spectrum analysis method for NaI-type whole-body counter, Tsinghua University, 2018.

[40] R. Zhao, L.Y. Liu, J.L. Li, Continuum Subtraction Method under Full-energy Peak in γ Spectrum Based on Recursion Formula. At. Energy. Sci. and Technol. **52**, 1085 (2018). doi:10.7538/yzk.2017.youxian.0543

[41] D.V. Gopinath, K. Gopala, Analytical computation of the Compton continuum in gamma-ray spectrometry. Radiat. Phys. Chem. **56**, 525–534 (1999). doi:10.1016/S0969-806X(99)00342-4